The Think Aloud Method for the Validation of Program Evaluation Instruments with English and Spanish Speaking Youth

Michael C. Rodriguez & Maria Pabón
University of Minnesota

David A. Tafur
University of Illinois, Chicago

Nydia de Alba-Johnson
Mexico

April 9, 2007

Paper presented at the annual meeting of the
American Educational Research Association, Chicago, Illinois.

Citation:

Rodriguez, M.C., Pabón, M., Tafur, D.A., & Alba-Johnson, N. (2007, April). *The think aloud method for the validation of program evaluation instruments with English and Spanish speaking youth.* Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

The Think Aloud Method for the Validation of Program Evaluation Instruments with English and Spanish Speaking Youth

ABSTRACT

Program design to support positive youth development has been limited, in part because of the lack of research on the role of culture in positive Latino youth development and because of the limited evidence to support specific practices – most evaluations do not address the cultural aspects of youth development adequately. Four youth-development program evaluation instruments were administered to 25 native-Spanish and 25 monolingual-English speaking teens through the think-aloud methodology. Responses to each item were coded employing an asset-based coding system (coding associated deficits as needed). Responses were analyzed by item, across items within instruments, and across instruments. Results provided numerous inferences addressing the incomparability of instrument results given the substantial differences in intended meaning and understanding between the two groups.

Background

Assessment of linguistically and culturally diverse populations is challenging because of methodological, conceptual and practical difficulties during such research. However, demographic changes and the increasing number of individuals whose first language is other than English are forcing educational researchers and evaluators to realize the need for research focused specifically on these populations.

The 2000 census determined that over 20 million individuals living in the United States speak Spanish as their native language and approximately one in eight people identify themselves as Hispanic/Latino (Therrien and Ramirez, 2000). The US Census Bureau (2005) reported that over 31 million individuals living in the United States who are 5 years old or older speak Spanish as their native language and that approximately one in seven people identify themselves as Hispanic or Latino. Nearly one in five students enrolled in elementary and secondary schools in the US are Latino, three times the proportion from 1970. These numbers reflect the relevance to study the ways we evaluate youth development programs in minority language communities, especially Spanish Speaking communities. These numbers reflect the relevance to study and evaluate more carefully minority language communities such as the Spanish Speaking population. Evaluation instruments used with individuals who speak a language other than English may introduce processes that could alter the results of such evaluation. For instance language differences are generally associated with cultural differences that may influence the response patterns and interpretations given to concepts a tool is intended to measure (Okasaki and Sue, 1995). Consequently, special attention to issues related to language and culture is essential and needs to be taken into account when validating measurement or evaluation tools administered cross-culturally.

A method that allows researchers to clarify how individuals interpret and assign meaning to questions is the think aloud. The think aloud method requires individuals to verbalize as many of their thought processes as possible while responding to an instrument item (van Someren, Barnard & Sandberg, 1994). Think alouds are helpful in determining how participants understand the items they are reading to reveal information about their interpretation of text that is not always readily visible with other approaches (Jimenez, Garcia & Pearson, 1995).

Objectives

The main objective is to propose the use of the think aloud method as a means to validate the meaning of evaluation instruments that have been used to assess the impact of after-school and school-based programs with bilingual populations (Spanish/English); to investigate the extent to which items used to assess youth accurately reflect the characteristics and traits intended to be measured and whether the responses intended by youth are captured by the instrument. The arena of school-based and out-of-school time programs encompasses a significant portion of programs on the national agenda, promoting positive healthy youth development and school-related achievement. The evaluation of these programs is at the core of the national education and youth-related health agendas, culminating in the *What Works Clearinghouse* and evidence-based grant making at the federal level.

Theoretical Framework

Validity is the most fundamental consideration when developing and evaluating instruments. Validity has been defined "as the degree to which evidence and theory support the interpretations of the instrument's scores" (AERA, APA, NCME, 1999, p.9). Hence, validation of an instrument entails the accumulation of evidence that provides a strong basis for the interpretations proposed by the evaluator and involves careful attention to possible distortions in

meaning. In the assessment of individuals whose native language is other than English these distortions could arise as a consequence of linguistic and cultural differences (Okasaki, et al. 1995).

The youth development literature scarcely includes and reports on knowledge about Latino youth, and when it does, it is overwhelmingly deficit based in a narrow range of risk behaviors (Rodriguez & Morrobel, 2004). We currently do a poor job of evaluating Latino youth development programs with culturally relevant indicators and outcomes.

When youth development programs are evaluated, measurement and evaluation instruments are created or available standardized instruments are employed. In the context of Latino-based youth development programs, the instruments are administered to English and Spanish speaking youth and their responses are assumed to indicate the same message. To the degree that the intent in a student's response is a function of cultural differences (vis-à-vis language), the answer option may be the same but the intent or correct inference drawn from that answer option may be different. We would normally believe that similar responses from two groups of students means the students are saying the same thing – because the instrument is designed to measure a specific characteristic. However, because students of a different language background may rely on a different set of experiences, a different mind-set, decision process, or intentions in making their responses, the typical inferences drawn from the results will be invalid.

Because a measurement or evaluation instrument has been validated in one context does not allow us to generalize such validity to a different context. We do not validate instruments, we validate the inferences we draw from the scores of an instrument – we validate the uses of scores or decisions we make based on results of the measurement. We may be using youth development

evaluation instruments in cross-cultural contexts without the validity evidence to support decisions regarding the development, continuation, or closure of youth development programs.

In the assessment of individuals whose native language is other than English these distortions could arise as a consequence of cultural differences (Okasaki, et al. 1995).

Assessment of minority language populations can introduce construct-irrelevant components to the evaluation process where the interpretations and decisions made may be affected by cultural processes internal to the examinees. In the case of minority language populations these components might include cultural misunderstandings, emotional reactions, familiarity with conceptual terms, or language ability. In addition, cultural values can affect resulting interpretations by evaluating individuals from one culture with values appropriate to another culture (AERA, APA, NCME, 1999)

Studies in bilingualism and reading (Jimenez, Garcia & Pearson, 1995; Kamhi-Stein, 1998) have demonstrated that think alouds can reveal information about an individual's interpretation of text that is not always visible with other methods. In the same manner, the think aloud method can allow researchers to identify misunderstanding given to instrument items when culturally diverse populations are being evaluated; misunderstandings that are not easily identified when other sources of validity evidence are used.

Methods

Four representative program evaluation instruments used in after-school and school-based programs were selected for validation: (1) Piers-Harris Children's Self-Concept Scale (available in English and Spanish), (2) Quality of School Life Scale (English only), (3) ALAS program evaluation questionnaire (a California-based after school program, available in English and Spanish), and (4) Nolan's Decision Making and Problem Solving Scales (English only). These

instruments were submitted to the think aloud process with 25 native Spanish speakers (monolingual and bilingual) and 25 monolingual English speakers from middle and high schools in high poverty neighborhoods in Minneapolis, Minnesota.

Each student participated in a one-time think-aloud session one-on-one with a trained researcher. To begin the think-aloud session, we engaged each student in a brief training exercise to acquaint them with the think-aloud process, instructing them to complete two evaluation instruments (randomly assigned) while thinking aloud when responding to each item—a set of allowable prompts were designed a priori, including for example, "*Now what are you thinking?*" Prompts were designed to minimally interfere in the natural thinking process of the student (more on this in the paper). All think-aloud sessions were audio recorded (see Table 1 for process).

Table 1

Think-Aloud Procedures Followed by Interviewers

1. Introduce yourself
2. Turn recorder on – note time
3. Introduce study – read protocol instructions
4. Check for understanding (confirm consent)
5. Hand out instrument form and pencil
6. Complete think-aloud interview
7. Provide stipend
8. Thank student for participating

Responses to each item in each instrument were coded for content. The initial coding scheme was based on a list of 40 assets and associated deficits developed by the Search Institute (2006). Assets included areas such as support, empowerment, boundaries & expectations, constructive use of time, commitment to learning, positive values, social competence, and positive identity. Deficits were in related areas and also included issues related to substance use,

involvement in crime, and other anti-social behaviors. Researchers have been calling for a more asset oriented approach to research in Latino youth development rather than the deficit-oriented models (medical model) that overwhelm the literature (Rodriguez & Morrobel, 2004). As comments from participants were reviewed and did not fit directly into one of the codes provided on the asset or deficit lists, new codes were created. Each response was independently coded by two bilingual researchers; inconsistencies were reviewed by the two bilingual researcher and a third native English speaker with some familiarity with Spanish. The three researchers reviewed coding disagreements and resolved each through consensus – a process that strengthened the overall coding scheme and understanding of the approach to responding by each participant (more on this in the paper).

Code Development and Training

The initial list of codes included 40 assets from the Search Institute's (2006) *40 Developmental Assets for Adolescents*, and 24 deficits (Benson, 1993; Scales, 1996). A sample of items was identified to test out the coding scheme and process. When responses did not fit one of the codes from the original list of assets and deficits, a new response code was created and defined. The three project researchers coded items and met to discuss coding decisions. We also discussed the role of positive and negative comments and decided to allow each asset to serve either a positive function or negative function using the + or - signs in coding. For example, the first asset is "1. Family support", but if the comment was regarding the "lack of family support", the same code could be used with a - designator, essentially "-1."

Additional responses were temporarily coded to identify an exhaustive set of additional codes that were not included in the original list of assets and deficits. This generated nearly 50 additional codes, some of which were subcategories or more specific notions of original assets.

For example, one asset is "19. Religious community". A sub-asset consistent with this notion was "19.2 Religious beliefs influence actions."

The additional codes were developed independently as each researcher read and coded responses. The three independent lists of codes were then pooled by the lead researcher into a single set of common codes. The final code list consisted of the 40 original assets, 38 new assets, 24 original deficits and 8 new deficits, 110 codes in all. This final code list was then used to do operational coding.

Operational Coding

Each response was read and coded—one or more codes were ascribed to each response of each participant for each item. Two researchers coded each response and a third researcher resolved differences in codes. For the first 22 items of the Piers-Harris included in this report across the 110 possible codes, coder agreement was 72%. This agreement rate was based on 323 coded responses, where 72% was viewed as a strong agreement rate, given the fact that each of the 323 responses was subjected to a code list containing 110 codes.

Once a final list of applicable codes was resolved for each participant and responses to each item, the number of times a code was used in an instrument was tallied across participants and items. Because the number of items varied by instrument and the number of participants responding to each instrument varied to some degree, the number of times a code was used was then adjusted (normalized) to make the results comparable. The method used was done separately by language: the total count of times a code was used in an instrument was divided by the number of respondents and number of items then multiplied by 100 to facilitate interpretation (see Table 2). If every participant mentioned the same asset for every item, the resulting rate would be 100.

Differences in proportions were tested for statistical significance employing the independent samples *t*-test at $\alpha = .05$.

Table 2

Coding Analysis and Rate Computation

1. Presence of a code was tallied across items (*I* items) and participants (*J* participants) for a single instrument, for each asset/deficit (*k*), separately by language group (*l*):

   $$\sum_{i=1}^{I}\sum_{j=1}^{J} C_{kl} \ .$$

2. The total count for the presence of an asset/deficit was divided by the number of items n_i and the number of participants n_j for each language group (*l*), then multiplied by 100 to compute a rate for each asset/deficit and language group:

   $$rate_{kl} = \frac{\sum C_{kl}}{n_i n_j} \times 100$$

3. The rates for Spanish (S) participants were subtracted from the rates of English (E) participants to obtain a difference in rates of employing certain assets or deficits in their thinking:

   $Rate_{kE} - Rate_{kS}$ = *Difference in Rates for asset/deficit k.*

4. These rates and differences were computed for positive and negative codes separately, and a combined total.

Results

The main research question addressed by this study is: Do the evaluation instruments retain constant meaning when administered to native-English speaking versus native Spanish speaking youth?

We are currently completing analyses. The complete analysis will be a thorough accounting of responses within form (and in some cases by item) and then across forms. However, a significant pattern in response is emerging, based on the reference point for experiences, beliefs, or values that students draw on when responding to items. Spanish speaking students are much more likely to respond to an item from the perspective of family influence,

deferring to authority, or being inclusive in the reason for their particularly response – being "other" oriented. They are also more likely to not understand particular words or phrases and were more likely to skip questions entirely. In addition, the Spanish speaking students are more likely to say they had no experience with a particular item, including things like "allowance" and having parents that take away privileges such as staying out late or having access to their bicycle. Although English students also report to have limited access to some of these things (e.g., many reporting never having received an allowance) they knew what these things were.

A specific example of response differences includes issues related to language and comprehension. Spanish speaking youth were confused between the differences in similarly worded items. In the ALAS instrument, we find a typical question about "expected" and "desired" levels of educational achievement. Spanish speaking youth tended to give the same response because they thought the question was the same. The difference between what one expects versus desires to achieve is often interpreted as information about educational goals. A finding that Latino students have limited expectations about educational success would be unwarranted.

Some Spanish speaking youth reported that they needed to provide for their families, while others said that in order for them to go to college they needed to work (which some already reported doing). One student reported: "with a High School education, it should suffice to get a decent job." Others reported a desire to go to college for two years at most; this group recognized that it would help them achieve better job opportunities. These responses were quite distinct from the majority of native English speakers, who tended to report plans of going to college, law school, and graduate school. More specifically, Spanish speaking students had ideas of what they

would like to pursue as careers, but did not know what level of education would help them attain their goals. English speaking students did not have this problem.

The Piers-Harris Self-Concept Scale

In the Piers-Harris Self-Concept scale, we found Spanish speaking students uniquely commenting on issues related to culture, language, and racism. These responses introduce an element that is not directly assessed nor is part of the scoring of the instrument. Student responses address aspects of self-concept that are relevant to their thinking and personal experience, yet absent in interpretation of resulting scores. In fact, remarks from Spanish speaking students covered a wider range of topics (including 56 different coded comments) than those from English speaking students (including 32 different codes). The most common response coded for Spanish speaking students was "Does not understand the question", including 11.4% of all responses from Spanish speaking students. Other common responses for both English and Spanish speaking students can be found in Table 3.

Table 3

Most Common Responses on Piers-Harris by Language Group

English speaking students		Spanish speaking students	
Response	*%*	*Response*	*%*
33. Interpersonal competence	8.1	44. Do not understand the question	11.4
38.3 Emotional and academic strengths	6.6	21. Achievement motivation	7.4
38. Self-esteem	5.1	33. Interpersonal competence	6.5
45. Not giving up	4.5	12.3 Not causing trouble at school or home	5.7
38.1 Reports just enough self-esteem – generally positive	4.5	5. School climate	4.8
		45. Not giving up	4.5
		1. Family support	4.5
		37. Personal power	4.3

A couple of issues are important to note. First, Spanish speaking students were far less likely to understand particular statements (11.4% of responses) compared to English speaking students (2.5%). Both groups made frequent responses regarding Interpersonal Competence and Not Giving Up. English speaking students were more likely to refer to academic strengths (being smart) and self-esteem than were Spanish speaking students. Spanish speaking students were more likely to refer to being motivated to achieve, the importance of family support, and then Not Causing Trouble or School Climate issues. More than half of the responses to School Climate issues from Spanish speaking students were negative, referring to the role of an uncaring school climate in their responses (53% of the School Climate references were negative for Spanish speaking students). To investigate these differences in response rates between the two language groups, we computed the difference in response rates and report the largest differences below (see Tables 4 and 5).

Table 4

Differences that indicate more responses from Spanish speaking participants

Asset/Deficit	*Differences in rates*		
	Positive	Negative	Total
44. Do not understand	0.0	-8.8*	-8.8*
21. Achievement motivation	-5.2*	-1.1*	-6.4*
11. Family Boundaries	-2.6*	-0.3	-3.0*
5. Caring school climate	-0.8	-2.1*	-2.8*
12.3 Not causing trouble at school and home	-2.8*	0.2	-2.7*
30. Responsibility	-2.0*	-0.6	-2.6*
37. Personal power	-1.7	-0.6	-2.2*
41. Peer support	-0.9	-1.1*	-2.1*
56. Youth generally believes education is important	-2.0*	0.0	-2.0*
2. Parent communication	-1.7*	0.0	-1.7*
12. School boundaries are clear	-1.7*	0.0	-1.7*
25. Language – Limited English	-1.7*	0.0	-1.7*
30. Personal and family trouble and worries	-1.7*	0.0	-1.7*
1. Family support	-1.2	-0.3	-1.5
22. School engagement	-0.9*	-0.6	-1.4*
10.1 Feeling materialistic safety	-1.1*	0.0	-1.1*
39. Sense of purpose	-1.1*	0.0	-1.1*
26. Culture and cultural differences	-1.1*	0.0	-1.1*
28. Uncertainty	-1.1*	0.0	-1.1*
12.2 Understands the rules & limits among friends	-0.9*	0.0	-0.9*
21.2 Youth motivation comes from family effort to help	-0.9*	0.0	-0.9*
50. Modesty, acknowledges others, does not take all the credit	-0.9*	0.0	-0.9*
27. Racism	-0.9*	0.0	-0.9*
29. School difficulties	-0.9*	0.0	-0.9*
12.1 Understands the rules & consequences at school	1.2	-1.4*	-0.3

Note. Negatives indicate a higher rate for Spanish speaking students (English rate – Spanish rate).
* $p<.05$.

Aside from the large difference in Do Not Understand the Statement, we found significant differences among 25 response codes where Spanish speaking students were more likely to use the response code. For example, Achievement Motivation was more likely referred to by 6.4% more of the time among Spanish speaking students than English speaking students. Again, this

rate is out of 550 possible responses [at the time of this report, only 22 of the 40 items of the Piers-Harris had been included in the analyses: 22 items × 25 participants]. Others were large differences were found (at least 2.5% difference) included the role of Family Boundaries, School Climate, Not Causing Trouble, and sense of Responsibility. One response did not provide consistent results in both the positive and negative use: English speaking students were more likely (1.2% more) to refer to Understanding the Rules and Consequences at School in the positive sense, whereas Spanish speaking students were more likely (1.4% more) to refer to this in the negative sense, or suggesting that they did not Understand the Rules and Consequences at School.

Responses that were uniquely Spanish speaking student responses included issues related to parent communication, language and limited English proficiency, feeling materialistic safety (we have what we need to survive), and culture and cultural differences. It is important to note that the English speaking students represented the students found in the district, whom were about 1/3 White and 2/3 racial and ethnic minority students (including African America, Asian, and American Indian). Although only mentioned 4 times during the Piers-Harris interviews, comments regarding racism were only made by Spanish speaking students.

Table 5

Differences that indicate more responses from English speaking participants

Asset/Deficit	*Differences in rates*		
	Positive	Negative	Total
38.3 Emotional and academic strengths	0.2	2.7*	2.9
33.2 Very outgoing, loud, outspoken	2.5*	-0.3	2.2*
33.1 Youth chooses to be selective about interpersonal skills	0.0	2.2*	2.2*
37.1 Some things that are out of our control	0.9	0.7	1.7
38.1 Reports just enough self-esteem – generally positive	1.7	0.0	1.7
42. Does the work supposed to do, works hard, determined	1.5	0.0	1.5
33. Interpersonal competence	0.1	1.5	1.5
38. Self-esteem	1.8	-0.4	1.4
40. Positive view of personal future	1.0	0.0	1.0
16. High expectations	1.0	0.0	1.0
11.1 Understands the rules & consequences at home	-0.3	1.0	0.7
43. Generally positive mood	1.2	-0.6	0.6
17. Creative Activities	-0.5	1.0	0.5

Note. Positives indicate a higher rate for English speaking students (English rate – Spanish rate).
* $p<.05$.

Fewer response codes were more likely employed by English speaking students, largely due to the fact that their responses were more homogenous (employing a smaller variety of responses). We found significant differences among only 3 response codes where English speaking students were more likely to use the response. For example, Academic strength was more likely referred to by 2.9% more of the time (mostly in the negative sense) among English speaking students than Spanish speaking students. Others where significant differences were found included being Loud and Outspoken and being Selective about Interpersonal Skills (generally negative). However, there were some inconsistencies in rate differences based on whether the reference was positive or negative (these are shaded in the table). References to being Outgoing and Loud, having Self-Esteem, and having a Generally Positive Mood were more likely to be used in the positive sense by English speakers and more likely to be used in the

negative sense by Spanish speakers. The opposite was true for Understanding the Rules & Consequences and being involved in Creative Activities.

Educational Implications

Educational research has paid little attention to the relevance of cultural issues in the area of measurement leading frequently to misdiagnosis, overestimation or underestimation of results. The importance of improved evaluation with language and cultural minority populations is clear given the increasing number of new immigrants who are coming to reside in the United States. Therefore, investigating new validation processes for instruments assessing individuals of different backgrounds can result in a more adequate and fair evaluation process. In addition, validation of meaning of instruments used with culturally diverse populations such as the Spanish speaking population may help educational researchers to understand and learn cultural differences and values that should be taken into consideration when evaluating these populations. Finally students from different cultural backgrounds can benefit from assessment that makes an attempt to reflect their real opinions and ideas.

Accomplishing the task of evaluating cultural and linguistically diverse populations is a challenge that educational researchers increasingly encounter. The use of the think aloud method and/or focus groups as means to validate the meaning of evaluation instruments provides an opportunity to clarify the use of evaluation instruments and assessments on diverse populations.

Considerations for Future Think-Aloud Research with Youth Development Instruments

It seemed that the students looked for a “moral” in the questionnaire, several of them said “I think that this questionnaire/you are trying to tell us to study, to be good.” It appeared that they

looked for the “logic” behind the questions, and for the right/wrong answer. This speaks to the “face validity” of the instruments and students trying to figure out what the instrument was attempting to uncover.

A very important topic seemed to be education. From student responses, education appeared to be the way to get a better life than one’s parents, and some students even referred to a “promise” to the parents to study. Also, for several students, making one’s parents proud was an important reason to study.

Culture also seemed to be an important topic. Many students did not know what “ethnic group” meant, but appeared to have a construction of the concept of race as something that distinguishes people. It would be interesting to trace these conceptions of race across the interviews.

References

AERA, APA, NCME (1999). *Standards for educational and psychological testing.* Washington DC: American Education Research Association.

Benson, P.L. (1993). *The troubled journey: A portrait of 6th-12th grade youth*. Minneapolis, MN: Search Institute.

Jimenez, R.T., Garcia, G.E., & Pearson, P.D. (1995). Three children, two languages, and strategic reading: Case studies in bilingual/monolingual reading. *American Educational Research Journal, 32* (1), 67-97.

Kamhi-Stein, L.D. (1998). Profiles of underprepared second-language readers. *Journal of Adolescent & Adult Literacy, 41*(8), 610-619.

Okasaki, S., & Sue, S. (1995). Methodological issues in assessment research with ethnic minorities. *Psychological Assessment, 7*(3), 367-375.

Rodriguez, M.C., & Morrobel, D. (2004). A review of Latino youth development research and a call for an asset orientation. *Hispanic Journal of Behavioral Sciences, 26*(2), 107-127.

Scales, P.C. (1996). *Boxed in and bored. How middle schools continue to fail young adolescents—and what good middle schools do right*. Minneapolis, MN: Search Institute.

Search Institute. (2006). *40 developmental assets for adolescents (ages 12-18).* Minneapolis, MN: Author. Retrieved online June 20, 2006 at http://www.search-institute.org/assets/40AssetsList.pdf

Therrien, M, & Ramírez, R. R. (2000, March). *The Hispanic population in the United States.* U.S. Census Bureau, Current Population Report.

US Census Bureau (2005). *Facts for Features: Hispanic Heritage Month, 2005*. Retrieved online at http://www.census.gov/Press-Release/www/releases/archives/facts_for_features_special_editions/005338.html

Van Someren, M., Barnard, Y. F., & Sandberg, J.A. (1994). *The think aloud method: A practical guide to model cognitive processes.* London: Academic Press.

Asset and Deficit Codes

USES		
Positive	Negative	Assets (presence or absence)
		Support
+1	-1	1. Family support [or lack of family support]
+2	-2	2. Parent communication [negative family communication]
+3	-3	3. Other adult resources
+4	-4	4. Caring neighbors [or uncaring neighborhood]
+5	-5	5. Caring [or Negative] school climate
+6	-6	6. Parent involvement in schooling
		Empowerment
+7	-7	7. Community values youth
+7.1	-7.1	Community helps youth during their development
+8	-8	8. Youth as resource
+9	-9	9. Service to others
+10	-10	10. Safety
+10.1	-10.1	Feeling materialistic safety
		Boundaries & Expectations
+11	-11	11. Family Boundaries [or boundaries are unclear]
+11.1	-11.1	Understands the rules & consequences at home
+12	-12	12. School boundaries [or boundaries are not clear]
+12.1	-12.1	Understands the rules & consequences at school
+12.2	-12.2	Understands the rules & limits among friends
+12.3	-12.3	Not causing trouble at school and home (generally positive)
+13	-13	13. Neighborhood Boundaries
+14	-14	14. Adult role models
+15	-15	15. Positive peer influence
+16	-16	16. High expectations [too high expectations or low expectations are both negative]
		Constructive Use of Time
+17	-17	17. Creative Activities [or lack of creative activities]
+17.1	-17.1	Music talent, theater talent
+18	-18	18. Youth Programs
+19	-19	19. Religious community
+19.2	-19.2	Religious beliefs influence actions
+20	-20	20. Time at home
+21	-21	21. Achievement motivation [or not motivated]
+21.2	-21.2	Youth motivation comes from family effort to help
+22	-22	22. School engagement
+23	-23	23. Homework
+24	-24	24. Bonding to school [or doesn't care about school]
+25	-25	25. Reading for Pleasure

		Positive Values
+26	-26	26. Caring
+26.1	-26.1	Places high value on having friends
+27	-27	27. Equality and Social Justice
+28	-28	28. Integrity
+29	-29	29. Honesty
+29.1	-29.1	Being hypocritical (not being hypocritical is positive)
+30	-30	30. Responsibility [or lack of sense of responsibility]
+31	-31	31. Restraint
		Social Competencies
+32	-32	32. Planning and Decision Making
	-32.1	Youth makes decisions when needed – at the time (with the flow) [this is generally negative]
+33	-33	33. Interpersonal competence
	-33.1	Youth chooses to be selective about interpersonal skills [this is generally negative]
+33.2	-33.2	Very outgoing, loud, outspoken
+34	-34	34. Cultural competence
+35	-35	35. Resistance skills
+36	-36	36. Peaceful conflict resolution
		Positive Identity
+37	-37	37. Personal power [or no personal power]
+37.1	-37.1	Youth understands there are some things that are out of our control
+38	-38	38. Self-esteem
+38.1		Reports just enough self-esteem – generally positive
+38.2	-38.2	Positive or negative body image
+38.3	-38.3	Emotional and academic strengths
+39	-39	39. Sense of purpose
+40	-40	40. Positive view of personal future [or worries about future]
+40.1	-40.1	Positive view of world [negative view of world]
+41	-41	Peer support [or negative or lack of peer support]
+42	-42	Does the work supposed to do, works hard. determined
+43	-43	Generally positive mood [or negative mood]
	-44	Do not understand, do not know (no entiendo) generally negative
+45	-45	Not giving up [or quitting]
+46	-46	Expresses him/herself freely
+47	-47	General positive behavior (school & home) [or negative]
+48	-48	Youth is aware of economic status, living situation (generally positive)
+49	-49	
+50	-50	Modesty, acknowledges others, does not take all the credit
+51	-51	Has a job – works for own things
+52	-52	Youth's Culture is important to self and family
+53	-53	Knows about his/her culture, wants to learn about own culture
+54	-54	Language: Parents value English

+54.1	-54.1	Parents prefer Spanish as primary language
+54.2	-54.2	Parents prefer bilingualism, no preference
+55	-55	Language: Youth values English
+55.1	-55.1	Youth values Spanish as primary language
+55.2	-55.2	Youth prefers bilingualism, no preference
+56	-56	Youth generally believes education is important

Negative	Deficits/Risks
D1	1. Alcohol use
D2	2. Driving/Riding and drinking
D3	3. Cigarette/tobacco use
D4	4. Other drug use
D5	5. Sexually active
D6	6. Non-use of contraceptives
D7	7. Depression/Suicide attempt
D8	8. Fighting
D9	9. Police trouble
D10	10. Theft
D11	11. Weapon use
D12	12. School absenteeism
D13	13. Drop-out
D14	14. Vandalism
D15	15. Eating disorder
D16	16. Alone at home
D17	17. Self-serving values
D18	18. TV overexposure
D19	19. Stress
D20	20. Physical abuse
D21	21. Sexual abuse
D22	22. Parental addiction
D23	23. Social isolation (lack of care, support)
D24	24. Negative Peer pressure
D25	Language – Limited English
D26	Culture and cultural differences
D27	Racism
D28	Uncertainty
D29	School difficulties
D30	Personal and family trouble and worries
D31	Easily distracted
D32	Gender bias – discrimination

Think Aloud Protocol for Interviewer

We are working on a project that will help us to learn more about after-school programs and how they can work with and better understand youth. This project was started because we don't really know if some after-school programs are really working. To understand if programs are working, we usually do something called an evaluation – we ask participants, like you, to answer questions on a survey or evaluation form. Sometimes these evaluation forms are very helpful because students that fill them out understand all of the questions and think about them in the same way. But sometimes, students from different backgrounds think about the questions in different ways – so their answers mean something different. For example, if we asked students about how important student council is, everyone might say it is very important. The problem is that you might think it is important for a different reason: like to help you get into college, or because you think student council can influence the way students treat each other in school. These are very different meanings of the importance of student council.

We would like you to complete one of the evaluation forms that are commonly used to evaluate after-school and school-based programs. While you complete the form, we would like you to think out loud as you answer each question on the form. We will help you to see how this works and do a few exercises so you can practice thinking out loud. When you think out loud, we will have a better idea of how you interpret each question – so we can understand what it means to you. We are especially interested in knowing how students who speak Spanish interpret evaluation questions compared to students who do not speak Spanish.

There are no right or wrong answers to any of the questions on the evaluation form. The evaluation form is one that students in other programs have answered when their program is being evaluated. After we do some practice exercises, I will help you through the beginning of the evaluation form and you can just continue answering questions at your own pace. There is no time limit. We believe this should take about 45 minutes. Your name will not be written on the form or provided to anyone else – this is completely confidential. You can choose to stop the meeting at any time. If you complete 75% of the interview, we will provide you with a stipend of $15 for your participation, as we notified you and your parents or guardians on the permission form.

Now I just want to make sure that you understand what we will do here and that you agree to participate. Can you tell me in your own words what you think the study is about? Can you tell me what your job is during this meeting? Do you know of any reason why you should not participate – can you see any risks to you because of participating in the think aloud interview? Do you have any questions about confidentiality or your rights to end the session if you need to?